WHAT IS CLAIMED IS:

1. A computer-implemented method for identifying correlated columns from database tables, comprising:

determining correlation attributes for a first column and a second column from one or more database tables, the correlation attributes describing for each column at least one of the column and content of the column;

comparing the correlation attributes from the first and second column;

identifying similarities between the first and second column on the basis of the comparison;

on the basis of the identified similarities, determining whether the first and second column are correlated; and

merging the first and second columns only if the columns are determined to be correlated.

2. The method of claim 1, wherein identifying the similarities comprises:

determining a correlation value indicating a degree of correlation between the first and the second column; and

determining whether the correlation value exceeds a predetermined threshold.

3. The method of claim 1, further comprising:

if it is determined that the first and second column are correlated, displaying an indication to a user that the first and second column can be merged; and

in response to user input, merging the first and second column into a single column.

4. The method of claim 1, wherein the first column is a column of a first database table and the second column is a column of a second database table, the method further comprising:

determining correlation attributes for N columns from the first database table and M columns from the second database table, where N and M are integers;

comparing the correlation attributes from each of the N columns with the correlation attributes from each of the M columns to identify similarities between the N and M columns; and

on the basis of the identified similarities, determining whether one or more of the N and M columns are correlated.

5. The method of claim 4, further comprising merging each of the one or more of the N and M columns determined to be correlated.

6. The method of claim 1, further comprising:

determining, from the one or more database tables, metadata describing characteristics of each column; and

wherein the correlation attributes are determined on the basis of the determined metadata.

7. The method of claim 6, wherein the determined metadata describes for each column an attribute of a data value in the column.

8. The method of claim 6, wherein the determined metadata describes for each column at least one of:

(i) a label;

(ii) a comment;

(iii) a constraint;

(iv) a trigger;

(v) a name;

(vi) a data type; and

(vii) a column length.

9. The method of claim 1, further comprising:

determining, from the one or more database tables, statistical parameters associated with each of the columns; and

wherein the correlation attributes are determined on the basis of the determined statistical parameters.

10. The method of claim 9, wherein the determined statistical parameters describe for each column at least one of:

(i) a minimum value;

(ii) a maximum value;

(iii) an average value; and

(iv) a range of values.

11. The method of claim 1, further comprising:

determining, from the one or more database tables, ontological properties describing cognitive qualities associated with each column; and

wherein the correlation attributes are determined on the basis of the determined ontological properties.

12. The method of claim 11, wherein the determined ontological properties describe for each column at least one of:

(i) a synonym;

(ii) a parent node; and

(iii) an ancestor node.

13. The method of claim 11, further comprising:

determining, from the one or more database tables, metadata describing the ontological properties.

14. The method of claim 1, further comprising:

determining, from the one or more database tables, measurement units associated with each column; and

wherein the correlation attributes are determined on the basis of the determined measurement units.

15. The method of claim 14, further comprising:

determining, from the one or more database tables, metadata describing the measurement units.

16. The method of claim 14, wherein identifying the similarities comprises:

determining whether the first and second column are associated with similar measurement units.

17. A computer-implemented method for identifying correlated columns from database tables, comprising:

determining metadata for at least two columns from one or more database tables, the metadata describing characteristics of each column;

analyzing content from the at least two columns from the one or more database tables; and

determining a degree of correlation between the at least two columns using the determined metadata and the analyzed content.

18. The method of claim 17, wherein determining the degree of correlation comprises:

assigning a first correlation value to the determined metadata;

assigning a second correlation value to the analyzed content, wherein the first and second correlation values are different; and

calculating a total correlation value on the basis of the first and second correlation values.

19. The method of claim 18, further comprising:

merging the at least two columns if the total correlation value exceeds a predetermined threshold value.

20. The method of claim 17, wherein analyzing the content comprises

determining statistical parameters from the content of each column.

21. The method of claim 17, further comprising:

merging the first and the at least one second column if it is determined that the first and at least one second column are correlated.

22. A computer readable medium containing a program which, when executed, performs a process for identifying correlated columns from database tables, the process comprising:

determining correlation attributes for a first column and a second column from one or more database tables, the correlation attributes describing for each column at least one of the column and content of the column;

comparing the correlation attributes from the first and second column;

identifying similarities between the first and second column on the basis of the comparison;

on the basis of the identified similarities, determining whether the first and second column are correlated; and

merging the first and second columns only if the columns are determined to be correlated

23. The computer readable medium of claim 22, wherein identifying the similarities comprises:

determining a correlation value indicating a degree of correlation between the first and the second column; and

determining whether the correlation value exceeds a predetermined threshold.

24. The computer readable medium of claim 22, wherein the process further comprises:

if it is determined that the first and second column are correlated, displaying an indication to a user that the first and second column can be merged; and

in response to user input, merging the first and second column into a single column.

25. The computer readable medium of claim 22, wherein the first column is a column of a first database table and the second column is a column of a second database table, the process further comprising:

determining correlation attributes for N columns from the first database table and M columns from the second database table, where N and M are integers;

comparing the correlation attributes from each of the N columns with the correlation attributes from each of the M columns to identify similarities between the N and M columns; and

on the basis of the identified similarities, determining whether one or more of the N and M columns are correlated.

26. The computer readable medium of claim 25, wherein the process further comprises:

merging each of the one or more of the N and M columns determined to be correlated.

27. The computer readable medium of claim 22, wherein the process further comprises:

determining, from the one or more database tables, metadata describing characteristics of each column; and

wherein the correlation attributes are determined on the basis of the determined metadata.

28. The computer readable medium of claim 27,wherein the determined metadata describes for each column an attribute of a data value in the column.

29. The computer readable medium of claim 27, wherein the determined metadata describes for each column at least one of:

(i) a label;

(ii) a comment;

(iii) a constraint;

(iv) a trigger;

(v) a name;

(vi) a data type; and

(vii) a column length.

30. The computer readable medium of claim 22, wherein the process further comprises:

determining, from the one or more database tables, statistical parameters associated with each of the columns; and

wherein the correlation attributes are determined on the basis of the determined statistical parameters.

31. The computer readable medium of claim 30, wherein the determined statistical parameters describe for each column at least one of:

(i) a minimum value;

(ii) a maximum value;

(iii) an average value; and

(iv) a range of values.

32. The computer readable medium of claim 22, wherein the process further comprises:

determining, from the one or more database tables, ontological properties describing cognitive qualities associated with each column; and

wherein the correlation attributes are determined on the basis of the determined ontological properties.

33. The computer readable medium of claim 32, wherein the determined ontological properties describe for each column at least one of:

(i) a synonym;

(ii) a parent node; and

(iii) an ancestor node.

34. The computer readable medium of claim 32, wherein the process further comprises:

determining, from the one or more database tables, metadata describing the ontological properties.

35. The computer readable medium of claim 22, wherein the process further comprises:

determining, from the one or more database tables, measurement units associated with each column; and

wherein the correlation attributes are determined on the basis of the determined measurement units.

36. The computer readable medium of claim 35, wherein the process further comprises:

determining, from the one or more database tables, metadata describing the measurement units.

37. The computer readable medium of claim 35, wherein identifying the similarities comprises:

determining whether the first and second column are associated with similar measurement units.

38. A computer readable medium containing a program which, when executed, performs a process for identifying correlated columns from database tables, the process comprising:

determining metadata for at least two columns from one or more database tables, the metadata describing characteristics of each column;

analyzing content from the at least two columns from the one or more database tables; and

determining a degree of correlation between the at least two columns using the determined metadata and the analyzed content.

39. The computer readable medium of claim 38, wherein determining the degree of correlation comprises:

assigning a first correlation value to the determined metadata;

assigning a second correlation value to the analyzed content, wherein the first and second correlation values are different; and

calculating a total correlation value on the basis of the first and second correlation values.

40. The computer readable medium of claim 39, wherein the process further comprises:

merging the at least two columns if the total correlation value exceeds a predetermined threshold value.

41. The computer readable medium of claim 38, wherein analyzing the content comprises:

determining statistical parameters from the content of each column.

42. The computer readable medium of claim 38, wherein the process further comprises:

merging the first and the at least one second column if it is determined that the first and at least one second column are correlated.

43. A data processing system comprising:

at least one database having one or more database tables; and

a correlation manager for identifying correlated columns from the one or more database tables, the correlation manager being configured for:

determining correlation attributes for a first column and a second column from the one or more database tables, the correlation attributes describing for each column at least one of the column and content of the column;

comparing the correlation attributes from the first and second column;

identifying similarities between the first and second column on the basis of the comparison;

on the basis of the identified similarities, determining whether the first and second column are correlated; and

merging the first and second columns only if the columns are determined to be correlated.

44. A data processing system comprising:

at least one database having one or more database tables; and

a correlation manager for identifying correlated columns from the one or more database tables, the correlation manager being configured for:

determining metadata for at least two columns from the one or more database tables, the metadata describing characteristics of each column;

analyzing content from the at least two columns from the one or more database tables; and

determining a degree of correlation between the at least two columns using the determined metadata and the analyzed content.